Centering Resonance Analysis: A superior data mining algorithm for textual data streams

Award Information
Agency:
Department of Defense

Branch
Air Force
Amount:
$749,277.00

Award Year:
2005
Program:
STTR

Phase:
Phase II
Contract:
FA9550-05-C-0023

Agency Tracking Number:
F033-0063
Solicitation Year:
2003

Solicitation Topic Code:
AF03T011
Solicitation Number:
n/a
Small Business Information
CRAWDAD TECHNOLOGIES, L.L.C.
5412 W. Harrison Ct., Chandler, AZ, 85226
Hubzone Owned:
N

Socially and Economically Disadvantaged:
N
Woman Owned:
N

Duns:
129119561

Principal Investigator:
Dan Ballard
PI
(480) 615-0159
ballard-d@cox.net

Business Contact:
Kevin Dooley
COO
(480) 236-9588
dooley@crawdadtech.com

Research Institution:
ARIZONA STATE UNIV.
Meghan McKendry
ORSPA, PO Box 873503
Tempe, AZ, 85287
(480) 965-1436
Nonprofit college or university

Abstract
An increasing amount of information comes in the form of streaming text: news media, email, and even human conversation. Creating insight from streaming text has been challenging though, because current knowledge discovery systems depend on expert analysts to move from simple data to actionable knowledge. In order for such systems to be used outside a small cadre of experts, they must incorporate deep analytics, which are information rich and diagnostic in nature, and when combined with domain knowledge, point to action. In Phase I, Crawdad Technologies demonstrated the superior text mining performance and scalability of its Centering Resonance Analysis (CRA). A topic tracking system based on CRA had up to five times better precision than ones based on existing technologies. In Phase II, Crawdad will develop deep analytics that capture (a) the temporal, dynamic nature of streaming text, and (b) the social dynamic that underlies many streaming text applications. We will create Dynamic CRA (DCRA) and Multi-Agent DCRA to model streaming text in applications such as communication monitoring, news streams, team and group conversations, and customer service interactions. The use of DCRA and Multi-Agent DCRA will be tested in an actual news streaming environment.

* information listed above is at the time of submission.